42390.P10466 *Patent*

UNITED STATES PATENT APPLICATION

FOR

Geographic Location Determination Including Inspection Of Network Address

INVENTORS:

Kingsum Chow
Sumeet Dua
Brian Tjaden

Prepared by

Steven D. Yates
Reg. No. 42,242
(503) 264-6589

Express Mail mailing label number: EL034437546US

Geographic Location Determination
Including Inspection Of Network Address

Field of the Invention

The invention generally relates to geographic location determination, and more particularly to inspecting a network address to identify a geographic location.

Background

It is advantageous for a server to determine the geographic location of incoming client connections. For example, location determination facilitates media distribution compliance, such as for honoring sporting event black out requirements. Location determination also facilitates providing geographic sensitive advertising, sales offers, discounts, data stream sources, and the like, as well as client tracking and evaluation.

Typically, a web site identifies an incoming client Transmission Control Protocol / Internet Protocol (TCP/IP) address (hereafter IP address), and performs a reverse Domain Name Service (DNS) lookup to obtain a text name for the IP address. This text name can then be inspected to guess a geographic location. For example, a specific IP address may resolve to "cs.sfu.ca", from which can be deduced that the client is connecting from the Simon Frazier University in Canada. However, a problem with this technique is that many domain names cannot reliably be looked up. For example, popular ".to" and ".tv" domains indicate, from a reverse DNS lookup, that clients are respectively geographically located in Tonga and Tuvalu, notwithstanding their actually being based in the United States or another country.

Another technique is to inspect the "whois" domain name registry database to obtain registration details for a domain name. However, there the information within the database is arbitrary, and therefore it also cannot be relied upon. Thus, what is needed is a more reliable way to perform geographic location determination.

Brief Description Of The Drawings

The features and advantages of the present invention will become apparent from the following detailed description of the present invention in which:

FIG. 1 illustrates one embodiment for determining geographic locations for a client that connects to a server.

FIG. 2 illustrates an exemplary format for encoding a numeric network address with a text based network addresses.

FIG. 3 illustrates one embodiment for reporting estimated client geographic locations.

FIG. 4 illustrates a suitable computing environment in which certain aspects of the invention may be implemented.

Detailed Description

As will be discussed below, many network addresses are named assigned a text based "human readable" address that is constructed with respect to known geographic locations, e.g., airports, cities, states, corporations, schools, etc. By inspecting geographic references in text based addresses assigned to routers and/or hosts

situated between a client and server on a network, a server may improve estimates of a geographic location for a client.

FIG. 1 illustrates one embodiment for determining a geographic location for a client that connects to a server. In the illustrated embodiment, multiple servers are configured to appear to the client as a single server **300** (FIG. 3).

A first operation is to receive **100** a client connection. Assumed an incoming client connection represents a connection by an individual computer, such as an end-user's computer. It will be appreciated, however, that the incoming connection may be from any networked device, e.g., mobile or non-mobile computers, phones, personal digital assistants (PDAs), etc.

The network address for the connecting client is then determined **102**. A network address represents a network identifier at which the incoming client may be reached. It is assumed that the client has a conventional numeric TCP/IP address, e.g., a dot quad address such as 192.168.10.100, or a text based network address. However, it will be appreciated that other network protocols may use a different addressing format.

A trace route is then performed **104** between the server and the client's network address. Trace routing involves determining a network path between the client and server. Examples of extant trace route programs include the "traceroute" application program provided by many Unix operating systems, and the "tracert.exe" application program provided by some Microsoft Windows operating systems. (Please note that all marks used herein are the property of their respective owners.)

In a TCP/IP network, trace routing is effected by directing towards the client successive network data packets with incrementally longer time-to-live (TTL) values. The TTL determines how many hops a packet is allowed before it is returned by a receiving host; return identifies the receiving host. Through successive TTL increments, all intermediary hosts (e.g., computers, routers, machines, other network devices, etc.) between the client and the server can be identified. It will be appreciated that other network environments may provide equivalent techniques.

In one embodiment, the trace routing is performed entirely externally to the client, e.g., the route path is always directed towards the client. In another embodiment, a trace route is performed from the client to the server. In a further embodiment, both the client-side and server-side trace routes are combined to maximize ability to determine a geographic location for the client. In one embodiment, the client-side trace route is performed by a network browser "plug in" or "helper application." In one embodiment, client side trace routing may be triggered automatically, such as by the client receiving and executing a server-side or client-side script, a web page, or other trigger.

Assuming trace routing may provide a results list comprising both text based and numeric network addresses, the numeric network address are lookup up **106** to determine their text based encoding. In one embodiment having a TCP/IP network, looking up network addresses comprises performing a reverse DNS lookup on the numeric network address. In one embodiment, rather than looking up all numeric addresses in the list, instead less than all are looked up. For example, one might only look up a few of the network addresses in the list "nearest" the client.

After text based encodings have been looked up, a format for the text based encodings is identified **108**. A format describes the arrangement, or structure, of the text assigned to a numeric network address. Frequently, network backbone companies structure the text based network addresses to facilitate organization, management and security of the network addresses. In particular, the structure often comprises a geographic component so that the network backbone can distinguish addresses assigned to different regions of the country.

For example, FIG. 2 illustrates an exemplary format utilized by UUNet (a division of WorldCom, Inc. of Georgia) for encoding a numeric network address with a text based network addresses. As illustrated, the exemplary network address has a first portion **200** comprising port and device data for network equipment utilized to host a particular network address, and a last portion **204** identifying a particular backbone provider (here, *alter.net* is a part of WorldCom). The middle portion **202** of the illustrated exemplary address comprises a reference to a nearest airport to the device to which the network address is assigned. Thus, one may look up the airport code for a UUNet address and determine the nearest airport for a particular network address.

Continuing with FIG. 1, after identifying **108** the format, a format description is looked up for the identified format. A format description identifies portions of the format, if any, that contain geographically significant data. For example, as illustrated in FIG. 2, the format description would identify the airport code portion **202** as being a geographically significant portion. Note that a text based network address may have

multiple geographic references, e.g., multiple airport, city, state, etc. identifiers. The format description identifies portions relevant to determining the geographic location of the network address.

In one embodiment, a database stores known formats and indicators of geographically significant portions of the stored formats. In one embodiment, if an address does not match any stored known formats, or if it appears to match multiple formats, then an expert system, rule based system, or other deductive system may be utilized to analyze a text based network address to determine its geographic location. For example, if a trace route indicates network traffic traveled from X, through Y, to Z, and it is determined X and Z are WorldCom addresses, but it is unclear what format Y has, then a rule may conclude that Y is also a WorldCom address by virtue of its being enclosed by WorldCom addresses. In one embodiment, a scoring system is used to select a most likely format for a particular network address.

Lexical analysis or pattern matching (e.g., regular expressions) may be used to match a text based network address against known formats. In one embodiment, the search space for a matching format is reduced by identifying the domain name of the network address, e.g., FIG 2. item **204**, and then only searching for matches used with that domain name. For example, if *alter.net* is identified in the last portion of the text based network address, then only formats used by *alter.net* are inspected for a match.

After looking up **110** the format description, geographically significant portions of the text based network address are extracted **112** and used to estimate a geographic location of the client. In the illustrated embodiment, rather than identifying the format and geographically significant portions of all network addresses resulting from the trace

route, instead only a network address "nearest" the client is processed. After extracting **112** the geographically significant portions, a test **114** is performed to determine whether refinement is desired. Such refinement may be required when no reverse DNS lookup can be performed to obtain a text based readable encoding of the "nearest" address, or if one desires to corroborate estimates through inspection of other addresses.

If no refinement is desired, then processing ends **116**. If refinement is desired, then as discussed above, a text based encoding for another network address is looked up **118**, if necessary, its format identified **120**, the format description looked up **122**, and geographically significant portions extracted **124**. This supplementary geographic location data is then used to revise **126** the initial geographic location estimate. Processing continues with another test **114** for further refinement.

In one embodiment, client geographic location estimation may also be based at least in part on data known about the client, e.g., from data obtained from client records, mailing lists, marketing research, etc. In one embodiment, a database is used to store text based encodings for trace route results that had to be looked up, as well as estimated geographic locations for client network addresses. This database may then operate as a cache for subsequent processing of repeated network addresses.

FIG. 3 illustrates one embodiment for reporting estimated client geographic locations. Estimated locations may be determined as discussed above, or retrieved from a database caching previous determinations.

As illustrated, a client may contact a single server, or multiple servers in a data center **300**. In one embodiment, data center servers may be logically grouped to appear as a single server. In one embodiment, the servers **300** host a customer's Internet web site(s). Contact activity, e.g., by customers, visitors, etc., results in the generation of client activity logs **302** containing network addresses associated with contacting entities. As discussed above, a trace route can be performed between the servers **300** and a client to estimate a geographic location for the client. In one embodiment, each client network address is immediately processed to identify a geographic location for the client upon the client contacting the servers **300**. In another embodiment, network addresses are collected for later asynchronous processing when a sufficient number of network addresses have been collected.

Assuming that addresses are collected for later group processing, in one embodiment, client activity logs **302** are filtered **304** to remove undesirable network addresses to prevent these network addresses from being processed. Undesirable network addresses include addresses that have been previously processed, as well as recognized addresses, such as ones belonging to machines of the servers **300**, or other known/undesirable machines. It will be appreciated that various filter characteristics may be used to determine undesirable addresses.

Filtering results in a list **306** of network addresses for which a geographic location is to be determined. A test **308** is performed to determine whether a particular network address in the list has previously been located. If not, then the network address is asynchronously trace routed **310** to identify, as discussed above, intermediaries between the client and the servers **300**. As illustrated, it is assumed the trace route

operation also analyzes the route results to estimate a geographic location for the network address as discussed above. The trace routing is performed asynchronously to allow collection of client network addresses for location to continue independent of the trace routing operation. It will be appreciated, however, that some embodiments may perform the trace routing synchronously, such as discussed above, when a client network addresses is processed on contact with the servers **300**.

The determined location for the client is then stored **312** in a database, and given to a data feeder **314** which is used to feed the results to a report generator **316**. The report generator generates reports **318** of client activity that can be distributed to businesses being hosted by the servers **300**. By storing **312** the determined location, in a subsequent geographic location determination, the test **308** returns that the client network address is already known, and the value stored **312** in the database is provided directly to the report generator **316**. It will be appreciated that various reports **318** may be generated, such as reports for a particular site hosted by the servers indicating the geographic location for clients contacting the sites. It will be further appreciated that the mechanisms discussed herein may be applied in real time determinations of appropriate advertising, content, etc. to be sent to a contacting client.

FIG. 4 and the following discussion are intended to provide a brief, general description of a suitable computing environment in which certain aspects of the illustrated invention may be implemented.

For example, an exemplary system for embodying one of the servers **300** of FIG. 3 includes a machine **400** having system bus **402** for coupling various machine

components. Typically, attached to the bus are processors **404**, a memory **406** (e.g., RAM, ROM), storage devices **408**, a video interface **410**, and input/output interface ports **412**. The machine may also include embedded controllers, Programmable Logic Devices (PLD), Programmable Logic Arrays (PLA), Programmable Array Logic (PAL), Generic Array Logic (GAL), Field-Programmable Gate Arrays (FPGA), Application Specific Integrated Circuits (ASIC), computers, smart cards, or another machine, system, etc.

The machine is expected to operate in a networked environment using logical connections to one or more remote machines **414**, **416** through a network interface **418**, modem **420**, or other communication pathway. Machines may be interconnected by way of a wired or wireless network **422** including an intranet, the Internet, local area networks, wide area networks, cellular, cable, laser, satellite, microwave, Blue Tooth, optical, infrared, or other carrier technology.

The invention may be described by reference to different high-level program modules and/or low-level hardware contexts that may be stored in memory **406** and/or storage devices **408**. Program modules include procedures, functions, programs, components, data structures, and the like, for performing particular tasks or implementing particular abstract data types. One skilled in the art will realize that program modules and low-level hardware contexts can be interchanged with low-level hardware instructions, and are collectively referenced hereafter as "directives." One will further appreciate that directives may be recorded or carried in a compressed, encrypted, or otherwise encoded format without departing from the scope of this patent,

even if the instructions must be decrypted, decompressed, compiled, interpreted, or otherwise manipulated prior to their execution or other utilization by the machine.

Memory **406**, storage devices **408**, and associated media, can store data and directives for the machine **400**. Program modules may be implemented within a single machine, or processed in a distributed network environment, and stored in both local and remote memory. Memory and storage devices include hard-drives, floppy-disks, optical storage, magnetic cassettes, tapes, flash memory cards, memory sticks, digital video disks, biological storage, and the like, as well as wired and wireless transmission environments, such as network **422**, over which directives may be delivered in the form of packets, serial data, parallel data, or other suitable transmission format.

Thus, for example, with respect to the illustrated embodiments, assuming machine **400** operates a server, then remote devices **414**, **416** may respectively be clients contacting the server over the network **422**. It will be appreciated that remote machines **414**, **416** may be configured like machine **400**, and therefore include many or all of the elements discussed for machine. It should also be appreciated that machines **400**, **414**, **416** may be embodied within a single device, or separate communicatively-coupled components, and may include or be embodied within routers, bridges, peer devices, web servers, etc.

Illustrated methods, and corresponding written descriptions thereof, are intended to illustrate machine-accessible media storing directives, or the like, which may be incorporated into single and multi-processor machines, portable computers, such as handheld devices including Personal Digital Assistants (PDAs), cellular telephones, and the like. Directives, when accessed, read, executed, loaded into, or otherwise utilized

by a machine, causes the machine to perform the illustrated methods. The figures, written description, and claims may variously be understood as representing instructions taken alone, instructions as organized in a particular form, e.g., packet, serial, parallel, etc., and/or instructions together with their storage or carrier media.

Having described and illustrated the principles of the invention with reference to illustrated embodiments, it will be recognized that the illustrated embodiments can be modified in arrangement and detail without departing from such principles.

And, even though the foregoing discussion has focused on particular embodiments, it is understood that other configurations are contemplated. In particular, even though expressions such as "in one embodiment," "in another embodiment," or the like are used herein, these phrases are meant to generally reference embodiment possibilities, and are not intended to limit the invention to particular embodiment configurations. As used herein, these terms may reference the same or different embodiments, and unless implicitly or expressly indicated otherwise, embodiments are combinable into other embodiments. Consequently, in view of the wide variety of permutations to the above-described embodiments, the detailed description is intended to be illustrative only, and should not be taken as limiting the scope of the invention.

What is claimed as the invention, therefore, is all such modifications as may come within the scope and spirit of the following claims and equivalents thereto.